NSIDC DAAC Workflows

DIF Workflow

1. NSIDC makes metadata in DIF format available for harvesting via an OAI-PMH server.
2. GCMD Science Coordinator harvests new/updated records via the OAI-PMH server twice a month.
3. GCMD Science Coordinator copies metadata files into DIF-Queue-In folder, which then copies the files to the Database Queue.
4. GCMD Science Coordinator performs QA/QC of metadata in Queue before committing to the database.
5. GCMD Science Coordinator contacts NSIDC staff informing them that their metadata has been published.

ECHO Workflow

NSIDC ECS:

1. Data providers stage data for the NSIDC ECS system to pick up via various methods (e.g., ftp, scp).
2. Data and metadata are ingested into the ECS system, consisting of databases, archive and ftp area.
3. The ECS Bulk Metadata Generation Tool (BMGT) picks up new or updated collection- and granule-level metadata, including URLs pointing to the NSIDC's ftp holdings, in near real time and exports them to ECHO.
4. ECHO ingests the metadata.
5. Access Control Lists (ACLs) are created by the NSIDC operator via ECHO's Provider User Management System (PUMP) in order to control visibility of each dataset and its holdings. The ACLs default to "hidden" so a dataset and its granules are only made publicly visible when the ACLs are set accordingly.
6. Periodically, automated and manual reconciliation processes are run in order to ensure that the ECS and ECHO holdings have not gotten out of sync. The reconciliations go through BMGT and corrections are made on either end according to a predefined set of rules.

NSIDC V0:

1. Data providers contact NSIDC about data that they would like to send. NSIDC operators either arrange an area for the data to be pushed to or they get the data themselves. In either method the data is placed in the staging area.
2. The writers are notified that there is new data and metadata/documentation is either created or updated. All of the metadata information is stored in the NSIDC Enterprise Database (EDB).
3. Data is then moved to the public ftp area and is flagged in the EDB as public.
4. NSIDC has a DAAC-Unique Extension(DUE) that is in place that periodically queries the EDB for new or updated collections.
5. The DUE will create xml files that are ftp'd to an ingest area at ECHO. This area is periodically checked and then ingested into the ECHO system.
6. Access Control Lists (ACLs) are created by the NSIDC operator via ECHO's Provider User Management System (PUMP) in order to control visibility of each dataset and its holdings. The ACLs default to "hidden" so a dataset and its granules are only made publicly visible when the ACLs are set accordingly.

EMS Workflow for NSIDC and NSIDCV0

1. Data provider generates the ASCII flat file in the format as defined in the ICD for the metadata of the data products that are being archived or distributed.
2. Data provider uploads this file to the EMS designated server and folder.
3. EMS processes the file and if there is any error EMS informs the data provider the error.
4. Data provider submits a revised file if original file had errors.
5. During the processing of metrics data if EMS notices product(s) for which EMS has no metadata (coming from Ingest, Archive, or searchExp files), EMS initiates a notification to the data provider for requesting the required metadata for the designated products(s). Data provider submits a new metadata or/and searchExp file that has complete metadata.EMS uses searchExp (containing regular expressions matched via perl script) file to assign product to the requestpath and lookup the metadata to the product. Information missing in one or the other file will cause incorrect metrics. Therefore, is important for the data provider to keep these two files current and in sync.
6. Going forward with CMR, any request for new metadata will be generated by the EMS and sent to CMR to gather metadata on the designated products from the data providers.

CMR Workflow

TBA

DIF/ECHO Metadata Contact

Deann Miller

Email: deann.miller@nsidc.org

EMS Contact:

Cathy Fowler
303-492-3397
cathyf@nsidc.org

Daniel Crumly
crumlyd@nsidc.org